

Final Report for Period: 09/1998 - 08/2002

Submitted on: 02/24/2003

Principal Investigator: Lee, Eva K.

Award ID: 9721402

Organization: GA Tech Res Corp - GIT

Title:

Investigations in Mixed Integer Programming

Project Participants

Senior Personnel

Name: Lee, Eva

Worked for more than 160 Hours: Yes

Contribution to Project:

Post-doc

Name: Easton, Todd

Worked for more than 160 Hours: No

Contribution to Project:

Todd Easton spent one year working with the PI while he was a Ph.D. student at Georgia Tech. His main focus was to understand the polyhedral theory arise from the MIP instances related to the PI's research projects. He graduated in July 1999, and is currently working with the PI on related projects in integer programming. There is no GRA nor postdoctoral support from this grant.

Graduate Student

Name: Brooks, J. Paul

Worked for more than 160 Hours: No

Contribution to Project:

J. Paul is a first-year Ph.D. student which has begun some preliminary study courses on integer programming with the PI. No GRA support is provided from this award.

Name: Campbell, Ann

Worked for more than 160 Hours: No

Contribution to Project:

Ann Campbell spent six months with the PI learning to apply mixed integer programming approaches to machine learning problems. She received no stipend support from this grant.

Undergraduate Student

Technician, Programmer

Other Participant

Research Experience for Undergraduates

Organizational Partners

Memorial Sloan Kettering Cancer Institut

1. Provide clinical expertise, clinical equipment and patient data for the clinical tests.

2. Collaborators on prostate cancer research

Marco Zaider, Ph.D., Attending Physicist, Head, Brachytherapy Physics, Memorial Sloan Kettering Cancer Institute; Professor, Clinical Radiation Oncology and Public Health, Department of Radiation Oncology, College of Physicians & Surgeons of Columbia University.

Peter Scardino, M.D., Attending Urologist, The New York and Presbyterian Hospital-Cornell Medical Center, Chief. Urology Service, Department of Surgery, Head, Prostate Cancer Program, and Murray F. Brennan Chair in Surgery, Memorial Sloan Kettering Cancer Institute.

Howard Ira Amols, Ph.D., Chief Clinical Physics.

Michael Zelefsky, M.D., Associate Professor, Department of Radiation Oncology, Brachytherapy Service.

Kristin Zakian, Ph.D., Assistant Attending Physicist, Department of Medical Physics.

Jason Koutcher, Ph.D., Attending Physicist, Department of Medical Physics.

Lawrence Schwartz, M.D., Associate Attending Radiologist, Memorial Sloan Kettering Cancer Institute, Associate Professor of Radiology, Cornell University Medical College.

Columbia University

1. Collaborator on the prostate cancer research --

C.S. Wu, Ph.D., Associate Professor, Department of Radiation Oncology, College of Physicians & Surgeons of Columbia University, Assistant Attending, Radiation Oncology Service, the New York Presbyterian Hospital.

2. Collaborator on the discriminant analysis research --

Richard J. Gallagher, Ph.D., Research Scientist, Department of Medical Informatics.

Emory University School of Medicine

Paula Vertino, Ph.D., Assistant Professor, Cancer Biologist and Biochemist, Winship Cancer Institute

Collaborative research on pattern recognition and classification techniques for predicting aberrant CpG island methylation in human cancer.

Other Collaborators or Contacts

Georgia Institute of Technology: Collaborative Research

1. Collaborators on integer programming --

George Nemhauser, Ph.D., A. Russell Chandler III Chair and Professor, Industrial and Systems Engineering.

Ellis Johnson, Ph.D., Coca-Cola Chair and Professor, Industrial and Systems Engineering.

Todd W. Easton, Ph.D., Postdoctoral Fellow, Industrial and Systems Engineering

2. Collaborators on discriminant analysis --

Mark Prausnitz, Ph.D., Assistant Professor, School of Chemical Engineering.

Cheng Zhu, Ph.D., Professor, School of Mechanical Engineering.

Activities and Findings

Research and Education Activities: (See PDF version submitted by PI at the end of the report)

1. Mixed Integer Programming Applied to Permanent Prostate Implants

This research concerns a particular type of radiation therapy, known as brachytherapy, that involves the placement of radioactive sources (seeds) either in tumors or near tumors. In this approach radiation is emitted outward and limited to short distances. Thus, unlike external beam radiotherapy, where radiation must traverse normal tissue in order to reach the tumor, brachytherapy is much more localized and therefore reduces radiation exposure to normal tissue while allowing an escalation in the radiation dose. However, the optimal placement and dosage of the radioactive seeds is a difficult and challenging problem. In collaboration with researchers at Memorial Sloan Kettering Cancer Institute and Columbia Presbyterian Medical Center, for the first time, a novel approach using mixed integer programming was explored. Variations of a mixed integer programming model were developed for determining optimal seed placement and dosage levels. Three main avenues were investigated: (1) the effectiveness of the MIP models and the clinical quality of the resulting treatment plans, (2) the development of new computational techniques for solving MIP instances arising from the medical domain, and (3) the use of the resulting optimization models and solver for clinical studies.

2. Machine Learning/Discriminant Analysis} This research focuses on the computationally intensive area of statistical classification, which has close connections with machine learning, neural networks and pattern recognition. In particular, we investigated a problem in statistical discriminant analysis which we initially modeled as a non-linear MIP. The novel aspects of our approach are that it allows a level of control on misclassification probabilities and allows entities to be classified into a reserved judgment region. Many variations of the basic model are possible, including variations without non-linearities and variations with only continuous variables. Simulation and cross validation tests were used to analyze the effectiveness of the various models in deriving good classification rules. In addition to experimenting with model variations, effort was also placed on developing effective computational techniques for solving the MIP problem instances. Also, in collaboration with researchers from the Georgia Tech bioengineering institute and Emory University Winship Cancer Institute, we have begun applying our classification algorithm to three biomedical problems: 1) prediction of ultrasonic cell disruption for drug delivery; 2) discriminant analysis of motility and morphology data from human lung carcinoma cells placed on a purified extracellular matrix protein; and 3) genomic pattern recognition, and prediction of aberrant CpG island methylation in human cancer. Our approaches to these three problems are briefly outlined below.

1) Prediction of Ultrasonic Cell Disruption for Drug Delivery

- . Increase skin permeability to allow absorption of macromolecules
- . Transiently disrupt viable cells, enabling exogenous material

- to enter without cell damage
 - . Challenge: ultrasound effects are very difficult to control
 - . Drug delivery and acoustic spectra are governed by ultrasound induced cavitation
 - . Apply discriminant analysis models to uncover 'predictive rule' of ultrasound-mediated disruption of red blood cells using acoustic spectrum and percentage of cell permeability recorded
 - . Facilitate the development of real-time feedback control device for ultrasound assisted drug delivery.
- 2) Discriminant analysis of motility and morphology data from human lung carcinoma cells placed on purified extracellular matrix protein
- . Register lung tumor activities in-vitro on various ECM protein bases under videomicroscopy
 - . Identify patterns and special features of activities of lung tumor cells and normal cells within these protein bases.
 - . Apply discriminant analysis models to determine classification rules to predict and discriminate behavior of tumor cells and normal cells
 - . Possible target treatment and new protein-base therapy for lung tumor
- 3) Genomic pattern recognition, and prediction of aberrant CpG island methylation in human cancer
- . Develop predictive rules for methylated status associated with TMS-1 in breast cancer
 - Potential for reactivating genes being silenced by reversing DNA methylation, thus providing an exciting molecular target for chemotherapeutic intervention
 - Develop novel treatment strategies aimed at blocking or reversing methylated status
 - Develop methylation markers for cancer prediction, treatment and prognosis
 - . Employ CpG islands obtained from RLGS as training set (contain methylated and unmethylated CpG islands of known status via experimental validation)
 - . Develop pattern recognition techniques to identify hidden sequence patterns and motifs that are discriminatory in nature.
 - . Using the discriminatory attributes, apply discriminant analysis models based on the training set to obtain predictive rules
 - . Validate the correctness of the predictive rules by testing on a collection of CpG islands with unknown status (status known to cancer biologists)

Findings: (See PDF version submitted by PI at the end of the report)

1. Mixed Integer Programming Applied to Permanent Prostate Implants

Results of clinical tests on a collection of prostate cancer patients are exciting. The proposed MIP approach appears to be quite promising, capable of producing clinically sound and high quality treatment plans. The plans provide better dose distribution to the prostate and lower dose to nearby critical healthy structures compared to those obtained via the current manual approach. This translates to better tumor control and lower complication rate for cancer patients. From the MIP standpoint, the resulting MIP instances proved to be computationally very challenging for existing commercial MIP solvers. The PI developed a general-purpose MIP solver, with a specialized heuristic procedure and branching scheme for this application. The specialized solver can produce good heuristic solutions in a matter of minutes. This real-time planning system opens up new research possibilities within the medical domain. One such possibility, which involves incorporating complex biological relationships and time-volume-dose information within the MIP model, has been investigated. The extremely difficult MIP instances arising from this application also prompted investigation of theoretical and computational techniques for solving general dense MIP instances. New theoretical results were obtained and submitted for publication in Mathematical Programming.

2. Discriminant Analysis and Classification Techniques

The results of the cross validation tests show that our MIP approach is competitive with previous (unconstrained) discrimination methods when applied to well-separated datasets, and significantly reduces the misclassification rate for datasets in which the groups are more mixed. The primary importance of this work lies in the fact that it is apparently the first time that a viable computational approach has been offered for constrained discriminant analysis that allows any number of a priori groups. The linear models without integer variables provide competitive results compared to the MIP models when appropriate weights are used to control the emphasis between correct classification and allocation to the reserved judgement region. Very exciting results were obtained for the three biomedical applications.

1) The results for predicting ultrasound cell disruption indicate that the variable combination consisting of ultrasound exposure time and acoustic signals measured at the driving frequency and its higher harmonics yields the best rule. The discriminant analysis methods used for deriving the prediction rules are broadly applicable, and could be used to develop prediction rules in other scenarios involving different cell types or tissues. These rules and the methods used to derive them could be used for real-time feedback about ultrasound's biological effects.

2) The lung tumor study offers observations which correlate with behaviors seen {in vivo} and suggest specific roles for the extracellular

matrix proteins and their integrin receptors in metastasis. Cell translocation in vitro has been associated with malignancy, as has an elongated phenotype and a rounded phenotype. The present work suggests that extracellular matrix proteins contribute in different ways to the malignancy of cancer cells, and that multiple malignant phenotypes exist. Also, the behavior of highly metastatic Calu-1 cells is sufficiently distinct from that of normal L-132 cells when placed on purified collagen IV to develop an accurate and quantitative model predicting metastatic potential.

3) Epigenetic silencing associated with aberrant methylation of promoter region CpG islands is one mechanism leading to loss of tumor suppressor function in human cancer. Using DNA pattern recognition we identified seven discriminatory novel patterns. When applying our discriminant analysis techniques, a classification function was derived based upon the frequency of these seven sequence patterns that was capable of discriminating methylation-prone from methylation-resistant CpG islands with 82% accuracy. The data indicate that CpG islands differ in their intrinsic susceptibility to de novo methylation, and suggest that the propensity for a CpG island to become aberrantly methylated can be predicted based on its sequence context.

Training and Development:

1. Two graduate students and a postdoctoral fellow were involved in research related to this project. The students were involved in the modeling and computational investigation in the research. The postdoctoral fellow was involved in investigation of polyhedral theory on related MIP instances.

2. Establishment of Center for Operations Research in Medicine.

The PI's main research focus is computational and optimization methods for medicine. With the funding from NSF and the Whitaker Foundation, the PI has established the Center for Operations Research in Medicine. The center is a collaborative education and research center established between the School of Industrial and Systems Engineering at Georgia Institute of Technology and medical researchers in different disciplines. The PI is the director of the center. The center mission is to foster interdisciplinary education and research efforts involving the development and application of sophisticated techniques from the field of operations research to problems in medicine. The PI has included the following resources on the website (<http://www.isye.gatech.edu/~evakylee/medcalor>) for students, educators, and researchers:

- .Faculty list
- .Affiliated medical faculty
- .Research projects
- .Graduate program
- .Pre-med program (for GT undergraduates premed advisees of the PI)
- .Media News reports and publications

Outreach Activities:

1. Student Mentoring/Advising: Junior/Senior/Graduate level --

The principal investigator includes lectures on her research projects in courses on integer programming and operations research. The students participate in class projects which come from parts of the PI's research.

2. Courses taught --

The principal investigator is responsible for teaching junior/senior level courses in operations research and engineering optimization, and Ph.D. level integer programming courses. Students learn theoretical and computational issues regarding integer programming. Real applications are used to demonstrate the complexity of real-world problems and the difficulty of developing good realistic mathematical models and effective computational strategies. In Spring 2000, the PI offered a new course in OR -- "Topics in Optimization -- Applications to Real-World Problems" for senior IE/OR students. Students were engaged in projects from treatment planning optimization using MIP approaches, and other real applications within the PI's research domain. There were numerous positive comments about THIs new course; and below are a few of them.

Quoted comments from student evaluation:

- . This was probably one of the best courses I have taken at Georgia Tech thus far. Dr. Lee really got me to think

more about what can actually be done with IE skills. Plus her sincere concern for the students really brought us to class, and got us involved. THANKS A LOT DR LEE!!!

. This class was very interesting to me and challenging. However, it was manageable and not overly demanding. I have learned a great deal about optimization and its real world applications. Lee is very interested in the success of her students and encourages all of us to come to her for help with class or even personal issues. She is very devoted to teaching and to the subject matter. Great teacher!!!

. This was one of my most interesting classes as I was finally able to apply classroom knowledge to a real-world problem.

. This course was very beneficial. There are not many IE classes where we actually use methods learned in class and apply them to real-world problems. More classes like this are needed.

The PI also was involved in the development of a new course for Georgia Tech -- Cancer Biology and Biotechnology. This course was developed in Summer 2002 to facilitate the multi-disciplinary nature of cancer research from biological scientists, engineers and scientific computing researchers. The PI was one of the core faculty in designing and teaching this course. In particular, she was responsible for the topics in "cancer bioinformatics and computing". This course is offered once every year and is taught collaboratively by a group of active cancer researchers at Georgia Tech. It is offered at both the undergraduate level (senior) and graduate level.

3. Invited Seminars, News Media Presentations, and High-School Recruiting --

The PI has organized numerous sessions of optimization in medicine in major mathematical programming and operations research conferences. She was also invited to present talks at biomedical engineering department seminars and mathematical programming and operations research conferences. This has enabled her to introduce to diverse audiences the idea of applying computational optimization and operations research techniques to important problems in cancer treatment research.

The PI has lectured to different types of audiences (government officials, news reporters and journalists (who report to the general public), scientific community outside the PI's domain of research, doctoral students at international conferences, premed undergraduates, and high school girls) on cancer research and the importance of a multi-disciplinary effort in advancing this research frontier.

Major Presentations:

- . Discriminant Analysis and Predictive Models in Medicine, Department of Cardiology, Emory University School of Medicine, Feb 2003.
- . Computational Techniques for Solving Dense Integer Programming Instances, INFORMS San Jose Nov 2002.
- . Facets of the Independent Set Polytope, Integer Programming Conference in honor of Egon Balas, Carnegie Mellon University, Jun 2002.
- . Cancer Bioinformatics: Cutting-edge Cancer Technologies at Georgia Tech and Emory University, May 13 2002.
- . Advances in Revolutionizing Early Detection and Treatment of Cancer, Georgia Tech Media Briefing on Biotechnology Issues, Apr 12 2002.
- . Generating Cutting Planes for Mixed Integer Programming Problems in a Parallel Distributed Memory Environment, Department of Computing and Software, McMaster University, Hamilton Canada Mar 2002.
- . MRS-guided dose-escalation treatment planning optimization for permanent prostate implants, Hamilton Regional Cancer Centre, McMaster University, Hamilton Canada Mar 2002.
- . Computational optimization techniques for brachytherapy, external beam and IMRT treatment planning, National Cancer Institute and National Science Foundation Sponsored Workshop on Operations Research and Radiation Therapy, Washington DC Feb 2002.

- . Optimization with multiple objectives, and its application to radiation therapy, National Cancer Institute and National Science Foundation Sponsored Workshop on Operations Research and Radiation Therapy, Washington DC Feb 2002.
- . Building a Strong Foundation for an Academic Career, INFORMS 2001 Doctoral Colloquium, Miami Nov 2001.
- . Computational Optimization Techniques for Treatment Planning of Prostate Permanent Implants, INFORMS Miami, Nov 2001.
- . Computational Optimization Methods for Treatment Planning in Radiation Therapy, Washington University Medical School, St. Louis Aug 2001.
- . An Innovative Computerized Treatment Planning System for Permanent Prostate Implants, Annual Meeting of the American Association for the Advancement of Science, San Francisco Feb 2001.
- . Mixed Integer Programming Approaches to Permanent Radioactive Seed Implants for Prostate Cancer Treatment, INFORMS Cincinnati, May 1999.

Journal Publications

- D. Silvern, E.K. Lee, R. Gallagher, L.G. Stabile, R.D. Ennis, C.R. Moorthy, and M. Zaider, "Treatment Planning for Permanent Prostate Implants: Genetic Algorithm versus Integer Programming", Medical & Biological Engineering & Computing, p. 850, vol. 35, (1997). Published
- E.K. Lee, R.J. Gallagher, D. Silvern C.S. Wu, and M. Zaider, "Treatment Planning for Brachytherapy: An Integer Programming Model, Two Computational Approaches and Experiments with Permanent Prostate Implant Planning", Physics in Medicine and Biology, p. 145, vol. 44 (1), (1999). Published
- E.K. Lee, R. Gallagher, D. Patterson, "A Linear Programming Approach to Discriminant Analysis with a Reserved Judgment Region", INFORMS Journal on Computing, p. , vol. , (1999). Accepted
- M. Zaider, M. Zelefsky, E.K. Lee, K. Zakian, H.A. Amols, M. Zelefsky, J. Dyke, J. Koutcher, "Treatment Planning for Prostate Implants Using MR Spectroscopy Imaging", International Journal of Radiation Oncology, Biology and Physics, p. 1085, vol. 47(4), (2000). Published
- C.S. Wu, R.D. Ennis, P.B. Schiff, E.K. Lee, M. Zaider, "Dosimetric and Volumetric Criteria for Selecting a Source Type and Source Activity in the Presence of Irregular Seed Placement in Permanent Prostate Implants", International Journal of Radiation Oncology, Biology and Physics, p. 815, vol. 47, (2000). Published
- E.K. Lee, R.J. Gallagher, M. Zaider, "Mixed Integer Programming Approaches to Treatment Planning for Brachytherapy -- Application to Permanent Prostate Implants", INFORMS -- Operations Research, p. 147, vol. 119, (2003). Accepted
- E.K. Lee, R.J. Gallagher, A. Campbell, M. Prausnitz, "Prediction of Ultrasound Cell Membrane Disruption Using Statistical Analysis and Machine Learning Techniques Applied to Acoustic Spectra", IEEE Transactions on Biomedical Engineering, p. , vol. , (2001). Accepted
- EK Lee, M Zaider, "Determining an Effective Planning Volume for Permanent Prostate Implants", International Journal on Radiation Oncology, Biology, Physics, p. 1197, vol. 49(4), (2001). Published
- EK Lee, "Generating Cutting Planes for Mixed Integer Programming Problems in a Parallel Distributed Memory Environment", INFORMS Journal on Computing, p. , vol. , (2001). Accepted
- T Easton, K Hooker, EK Lee, "Facets of the Independent Set Polytope", Mathematical Programming, p. , vol. , (2002). Accepted
- EK Lee, JP Brooks, Adele Wright, C Zhu, "Discriminant analysis of motility and morphology data from human lung carcinoma cells placed on purified extracellular matrix proteins", Annals of Biomedical Engineering, p. , vol. , (2002). Submitted

EK Lee, M Zaider, "Intra-Operative Dynamic Dose Optimization in Permanent Prostate Implants", International Journal of Radiation Oncology Biology Physics., p. , vol. , (2002). Accepted

FA Feltus, EK Lee, JF Costello, C Plass, PM Vertino, "Predicting Aberrant CpG Island Methylation in Human Cancer", Proceedings of the National Academy of Science, p. , vol. , (2002). Submitted

Books or Other One-time Publications

E.K. Lee, R.J. Gallagher and M. Zaider, "Planning Implants of Radionuclides for the Treatment of Prostate Cancer: An Application of Mixed Integer Programming", (1999). Feature Article, Published
Bibliography: Optima, Mathematical Programming Society Newsletter No. 61

E.K. Lee, "Branch-and-Bound Methods", (1999). Book, Published
Editor(s): Panos Pardalos and Mauricio Resende
Collection: Handbook of Applied Optimization
Bibliography: Oxford University Press

E.K. Lee and J. Michell, "Branch-and-Bound Methods for Integer Programming", (1999). Book, Published
Editor(s): Chris A. Floudas and Panos M. Pardalos
Collection: Encyclopedia of Optimization
Bibliography: Kluwer Academic Publishers, The Netherlands.

EK Lee, M Zaider, "Treatment planning for low dose rate and high dose rate brachytherapy", (2003). Book, Accepted
Editor(s): Dicker, Merrick, Gomella, Valicenti, Waterman
Bibliography: Basic and Advanced Techniques in Prostate Brachytherapy

EK Lee, M. Zaider, "Treatment planning optimization in brachytherapy", (2003). Book, Accepted
Collection: Handbook of Operations Research/Management Science Applications in Health Care
Bibliography: Kluwer Academic Publishers

EK Lee, "A Linear-Programming Based Parallel Cutting Plane Algorithm for Mixed Integer Programming Problems", (1999). Proceedings, Published
Bibliography: Proceeding for the Third Scandinavian Workshop on Linear Programming

Reuter reporter, "Computerized Treatment May Help Prostate Cancer, EK Lee", (2001). News media published article, Published
Collection: Medical news
Bibliography: New York Times

Medical reporter, "Pinpoint radiation targets cancer of the prostate, EK Lee", (2001). News media published article, Published
Collection: Medical news
Bibliography: London Times

Science reporter, "Real-time computer planning optimizes seed placement, EK Lee.", (2001). News media published article, Published
Bibliography: Urology Times

Web/Internet Site

URL(s):

1. <http://www.isye.gatech.edu/~evakylee>
2. <http://www.isye.gatech.edu/~evakylee/medicalor>

Description:

1. This personal site consists of research projects conducted and led by the Principal Investigator. Abstracts of refereed papers are included and acknowledgements to NSF can be found within abstracts whenever appropriate.

2. Establishment of Center for Operations Research in Medicine.

The PI's main research focus is computational and optimization methods for medicine. With the funding from NSF and the Whitaker Foundation, the PI has established the Center for Operations Research in Medicine. The center is a collaborative education and research center established between the School of Industrial and Systems Engineering at Georgia Institute of Technology and medical researchers in different disciplines. The PI is the director of the center. The center mission is to foster interdisciplinary education and research efforts involving the development and application of sophisticated techniques from the field of operations research to problems in medicine. The PI has included the following resources on the website (<http://www.isye.gatech.edu/~evakylee/medicalor>) for students, educators, and researchers:

- .Faculty list
- .Affiliated medical faculty
- .Research projects
- .Graduate program
- .Pre-med program (for GT undergraduates premed advisees of the PI)
- .Media News reports and publications

Other Specific Products

Product Type: Teaching aids

Product Description:

Book chapters on applying integer programming techniques to biomedical applications have been prepared. The articles are being used in undergraduate and graduate courses in integer programming to introduce students to new application and research areas for integer programming investigation.

Sharing Information:

Articles are available to students at Georgia Tech. They are also available to outside colleagues upon request.

Product Type: Other inventions

Product Description:

An invention disclosure and patent entitled "Improved Brachytherapy Treatment Planning Method and Apparatus" was filed on October 1999.

Sharing Information:

The invention is useful to the medical community in designing optimal treatment plans for brachytherapy radiation treatment. NSF support for the research was acknowledged.

Product Type: Center for Operations Research in Medicine

Product Description:

A research center, entitled "Center for Operations Research in Medicine" has been established. The center fosters interdisciplinary education and research efforts involving the development and application of sophisticated techniques from the field of operations research to problems in medicine. Currently the PI has 8 Ph.D. students from "Optimization," "Algorithms, Combinatorics, Optimization," and "Bioinformatics." programs at Georgia Tech. There are also 40 undergraduate premed engineering students affiliated to the center.

Sharing Information:

The center has a website where students and other researchers can access -- <http://www.isye.gatech.edu/~evakylee/medicalor>. NSF support is acknowledged.

Contributions

Contributions within Discipline:

Mixed integer programming (MIP) is one of the major subfields in the broad field of operations research. It provides an extremely powerful framework for modelling a huge variety of problems arising from real-world applications. It is therefore used extensively in both government and industry. In addition, many (perhaps, even most) problems encountered in operations research have some optimization component associated with them. Often the optimization technology that is most appropriate is MIP. In this research funded by NSF, the PI developed MIP optimization models to address optimization problems arising in brachytherapy treatment (radioactive seed implants) planning for prostate cancer and in statistical discriminant analysis.

1. The MIP instances arising from prostate cancer treatment planning prove to be very difficult for competitive commercial optimization software. Unlike the sparse MIP instances arising from other industrial applications, a prostate cancer instance involves a dense constraint matrix; and commercial optimization solvers simply do not handle dense systems well. Thus, one major focus of the research includes investigating effective computational strategies for handling dense MIP systems and polyhedral studies. This provides a new direction of investigation for the MIP research community. A paper entitled "Facets of Independent Set Polytopes" has been prepared and was accepted for publication in *Mathematical Programming*. In particular, small dense MIP instances intractable for existing competitive commercial software were solved to proven optimality using cutting planes derived from our theoretical study. Another, entitled "Computational Issues for Solving Dense MIP Instances and Its Application to MIP Instances Arising from Treatment Plan Optimization for Radiation Therapy" is in preparation for submission to a refereed journal.

2. The discriminant analysis model developed is the first viable computational model for constrained discriminant analysis that allows any number of a priori groups. Furthermore, the approach using linear programming variations offers polynomial-time solvable models for this computationally intensive problem. The models developed proved to be viable in various types of prediction/classification problems arising from medical domain.

The PI began investigating operations research applications in medicine in 1996 (and subsequently received NSF funding in 1997). Since then, the PI has influenced research activities in this area by speaking and organizing invited clusters and sessions on mathematical programming in medical applications at conferences, editing a volume of *Annals of Operations Research* dedicated to optimization in medicine, and co-organizing an NCI-NSF funded workshop on operations research applied to radiation therapy. The goal of the workshop was to assemble an interdisciplinary group of researchers to give direction and to identify important issues that have the potential to advance the state-of-the-art in radiation treatment planning.

Contributions to Other Disciplines:

The mathematical models and computational methods developed have proven to be very useful in treatment planning optimization for prostate cancer patients. Prostate cancer is the second most common cancer among American males, and it is responsible for approximately 200,000 deaths per year. Improvement in treatment translates to improved local tumor control, less normal tissue complication and better quality-of-life for patients. The work was presented at the annual meeting of the American Association for the Advancement of Science, San Francisco Feb 2001, and subsequently reported upon in the news media, including *New York Times* and *London Times*, highlighting the importance of computational advances to these medical problems.

The discriminant analysis research has been both exciting and satisfying, due in large part to 1) the flexibility of the models developed, 2) the fact that the models have been successfully applied to different biomedical problems. Such application has provided insight to ultrasonic cell disruption for drug delivery, cell morphology and motility in lung cancer, and genomic patterns and classification in methylated CpG islands in human cancer. Each of these are important cancer research projects that are actively being studied by researchers in the respective specialties. A paper reporting the classification study on breast cancer has been submitted to the *Proceedings of the National Academy of Science*.

Results of the funding have fostered interdisciplinary research collaboration between engineering and medical researchers in the integer programming work; and between operations researchers and biomedical engineers in the discriminant analysis work. In particular, in addition to the sponsored projects mentioned in this report, other interdisciplinary collaborative research projects have resulted through the interaction. Current prostate cancer projects with medical researchers which will benefit from the use of sophisticated computational and optimization techniques involve biological optimization, and effective planning volume study (which involves MIP instances with multi-stage constraints). The PI is also leading research in brain tumor radiosurgery treatment planning optimization using MIP techniques, intensity-modulated radiation therapy optimization, and automated tumor volume contouring using computational geometry and numerical analysis approaches. In classification and machine learning areas, the PI is involved in predictive models for heart disease patients and discrimination of breast cancer patients to various biological and environmental factors.

Contributions to Human Resource Development:

Through her teaching and student mentoring, the PI has exposed undergraduate and graduate students to the opportunity of applying operations research and optimization skills to complex biomedical problems, some of which are novel applications that have not previously been approached using sophisticated mathematical and engineering techniques.

Although the PI has only secured funding for one Ph.D. student for this project from NSF, many Ph.D. students have expressed interest and enthusiasm in pursuing thesis research in this area. Through partial support from other external funding opportunities, The PI was able to support 2 Ph.D. students and 1 postdoc in learning and working on this project.

The PI advises undergraduate engineering students and pre-med students. She meets with these students on a regular basis and discusses with them about medical research related to her areas of expertise. Some of these undergraduates work on reading courses with the PI to learn about biotechnology developments through the PI's work.

In addition to teaching courses on optimization and computational techniques in her own department, the PI is also involved in the development and team teaching of a new course at Georgia Tech entitled "Cancer Biology and Biotechnology". This course, developed in Summer 2002, emphasizes the multi-disciplinary nature of cancer research from the perspective of biological scientists and engineers, and scientific computing researchers. The PI is one of the core faculty involved in designing and teaching this course, and is responsible for the topic "cancer bioinformatics and computing". This course is to be offered once every year, and co-taught by a group of active cancer researchers at Georgia Tech. It is offered at both the undergraduate senior and graduate levels.

The PI has also been involved in high school recruiting and an educational development forum to address high school female students on applications of mathematics and computer science in medicine.

Contributions to Resources for Research and Education:

The PI is actively promoting medical research among the MIP community, and likewise, bringing awareness of optimization techniques and algorithmic approaches to the medical community. She was invited to serve as the guest editor for a special issue of Annals of Operations Research on "Optimization in Medicine". In addition, in International Symposium on Mathematical Programming 2000, the PI was the invited Cluster Chair for a new cluster "Medical Applications". The goal of this cluster was to heighten awareness among the MIP researchers of the challenging problems in medicine. The workshop, sponsored by National Cancer Institute and National Science Foundation, brought together researchers in three different domains to discuss the multi-disciplinary effort in radiotherapy treatment planning.

The PI was invited to write book chapters on branch-and-bound and parallel mixed integer programming for 'Encyclopedia of Optimization' (Ed. Panos Pardalos and C.A. Floudas, Kluwer Academic Publisher), and 'Handbook of Applied Optimization' (Ed. Panos Pardalos and Mauricio Resende, Oxford University Press). The Encyclopedia is intended for anyone working in the field of optimization and anyone using optimization methods, including undergraduate students. The Handbook is designed as the premier reference tool for the field of applied optimization, and has recently been named runner up (honorable mention) for the Association of American Publishers' (AAP) Outstanding Professional and Scholarly Titles of 2002 in the computer science category. In addition, the PI has been invited to write a textbook on "Optimization techniques for biomedical applications".

Contributions Beyond Science and Engineering:

The prostate cancer treatment planning optimization research is a first-of-its-kind computerized approach to improve the treatment planning procedure for brachytherapy. Preliminary clinical experiments suggest that the MIP-based treatment plans are far superior to those generated using the current manual approach, providing better coverage and conformity to the tumor area, improving dose distribution, and lowering irradiation of critical healthy tissue. In addition, fewer radioactive seeds and needles are used. All of these factors translate to better tumor control (thus a lower recurrence rate) and a decrease in normal tissue complication (thus, the quality of life of cancer patients will improve). The optimization engine developed to solve the MIP instances for the treatment models can return plans in a fraction of the time needed to generate a plan manually. Thus, the routine use of a system incorporating this technology can allow real-time planning during the implantation session, and thus eliminate the existing limitations and implant difficulties when pre-operative manual planning is used. This in turn will reduce labor costs; and hence, will help to reduce the medical expenses for both patients and insurers. Furthermore, this research assists in pushing the medical frontier by providing a valuable research tool for other research activities conducted in the cancer treatment domain. The work was presented at the annual meeting of the American Association for the Advancement of Science, San Francisco Feb 2001, and subsequently reported upon in the news media, including New York Times and London Times, highlighting the importance of computational advances to these medical problems.

Classification and predictive models have always been an integral component in the study of cancer prediction and detection. The models developed have advantages over others as they allow classification over different groups of entities, and they allow reserved judgement. This in

turn allows for multi-stage classification over the reserved judgement region, where different discriminatory attributes can be used at each stage. A paper reporting the classification on breast cancer has been submitted to the Proceedings of the National Academy of Science.

Categories for which nothing is reported:

1. Mixed Integer Programming Applied to Permanent Prostate Implants

This research concerns a particular type of radiation therapy, known as brachytherapy, that involves the placement of radioactive sources (seeds) either in tumors or near tumors. In this approach radiation is emitted outward and limited to short distances. Thus, unlike external beam radiotherapy, where radiation must traverse normal tissue in order to reach the tumor, brachytherapy is much more localized and therefore reduces radiation exposure to normal tissue while allowing an escalation in the radiation dose. However, the optimal placement and dosage of the radioactive seeds is a difficult and challenging problem. In collaboration with researchers at Memorial Sloan Kettering Cancer Institute and Columbia Presbyterian Medical Center, for the first time, a novel approach using mixed integer programming was explored. Variations of a mixed integer programming model were developed for determining optimal seed placement and dosage levels. Three main avenues were investigated: (1) the effectiveness of the MIP models and the clinical quality of the resulting treatment plans, (2) the development of new computational techniques for solving MIP instances arising from the medical domain, and (3) the use of the resulting optimization models and solver for clinical studies.

2. Machine Learning/Discriminant Analysis

This research focuses on the computationally intensive area of statistical classification, which has close connections with machine learning, neural networks and pattern recognition. In particular, we investigated a problem in statistical discriminant analysis which we initially modeled as a non-linear MIP. The novel aspects of our approach are that it allows a level of control on misclassification probabilities and allows entities to be classified into a reserved judgment region. Many variations of the basic model are possible, including variations without non-linearities and variations with only continuous variables. Simulation and cross validation tests were used to analyze the effectiveness of the various models in deriving good classification rules. In addition to experimenting with model variations, effort was also placed on developing effective computational techniques for solving the MIP problem instances. Also, in collaboration with researchers from the Georgia Tech bioengineering institute and Emory University Winship Cancer Institute, we have begun applying our classification algorithm to three biomedical problems: 1) prediction of ultrasonic cell disruption for drug delivery; 2) discriminant analysis of motility and morphology data from human lung carcinoma cells placed on a purified extracellular matrix protein; and 3) genomic pattern recognition, and prediction of aberrant CpG island methylation in human cancer. Our approaches to these three problems are briefly outlined below.

1). Prediction of Ultrasonic Cell Disruption for Drug Delivery

- . Increase skin permeability to allow absorption of macromolecules
- . Transiently disrupt viable cells, enabling exogenous material to enter without cell damage
- . Challenge: ultrasound effects are very difficult to control
- . Drug delivery and acoustic spectra are governed by ultrasound-induced cavitation
- . Apply discriminant analysis models to uncover "predictive rule" of ultrasound-mediated disruption of red blood cells using acoustic spectrum and percentage of cell permeability recorded
- . Facilitate the development of real-time feedback control device for ultrasound assisted drug delivery.

2). Discriminant analysis of motility and morphology data from human lung carcinoma cells placed on purified extracellular matrix protein

- . Register lung tumor activities in-vitro on various ECM protein bases under videomicroscopy
- . Identify patterns and special features of activities of lung tumor cells and normal cells within these protein bases.
- . Apply discriminant analysis models to determine classification rules to predict and discriminate behavior of tumor cells and normal cells
- . Possible target treatment and new protein-base therapy for lung tumor

3). Genomic pattern recognition, and prediction of aberrant CpG island methylation in human cancer

- . Develop predictive rules for methylated status associated with TMS-1 in breast cancer
 - Potential for reactivating genes being silenced by reversing DNA methylation, thus providing an exciting molecular target for chemotherapeutic intervention
 - Develop novel treatment strategies aimed at blocking or reversing methylated status
 - Develop methylation markers for cancer prediction, treatment and prognosis

- . Employ CpG islands obtained from RLGS as training set (contain methylated and unmethylated CpG islands of known status via experimental validation)
- . Develop pattern recognition techniques to identify hidden sequence patterns and motifs that are discriminatory in nature.
- . Using the discriminatory attributes, apply discriminant analysis models based on the training set to obtain predictive rules
- . Validate the correctness of the predictive rules by testing on a collection of CpG islands with unknown status (status known to cancer biologists)

1. Mixed Integer Programming Applied to Permanent Prostate Implants

Results of clinical tests on a collection of prostate cancer patients are exciting. The proposed MIP approach appears to be quite promising, capable of producing clinically sound and high quality treatment plans. The plans provide better dose distribution to the prostate and lower dose to nearby critical healthy structures compared to those obtained via the current manual approach. This translates to better tumor control and lower complication rate for cancer patients. From the MIP standpoint, the resulting MIP instances proved to be computationally very challenging for existing commercial MIP solvers. The PI developed a general-purpose MIP solver, with a specialized heuristic procedure and branching scheme for this application. The specialized solver can produce good heuristic solutions in a matter of minutes. This real-time planning system opens up new research possibilities within the medical domain. One such possibility, which involves incorporating complex biological relationships and time-volume-dose information within the MIP model, has been investigated. The extremely difficult MIP instances arising from this application also prompted investigation of theoretical and computational techniques for solving general dense MIP instances. New theoretical results were obtained and submitted for publication in Mathematical Programming.

2. Discriminant Analysis and Classification Techniques

The results of the cross validation tests show that our MIP approach is competitive with previous (unconstrained) discrimination methods when applied to well-separated datasets, and significantly reduces the misclassification rate for datasets in which the groups are more mixed. The primary importance of this work lies in the fact that it is apparently the first time that a viable computational approach has been offered for constrained discriminant analysis that allows any number of a priori groups. The linear models without integer variables provide competitive results compared to the MIP models when appropriate weights are used to control the emphasis between correct classification and allocation to the reserved judgement region. Very exciting results were obtained for the three biomedical applications.

1) The results for predicting ultrasound cell disruption indicate that the variable combination consisting of ultrasound exposure time and acoustic signals measured at the driving frequency and its higher harmonics yields the best rule. The discriminant analysis methods used for deriving the prediction rules are broadly applicable, and could be used to develop prediction rules in other scenarios involving different cell types or tissues. These rules and the methods used to derive them could be used for real-time feedback about ultrasound's biological effects.

2) The lung tumor study offers observations which correlate with behaviors seen *in vivo* and suggest specific roles for the extracellular matrix proteins and their integrin receptors in metastasis. Cell translocation *in vitro* has been associated with malignancy, as has an elongated phenotype and a rounded phenotype. The present work suggests that extracellular matrix proteins contribute in different ways to the malignancy of cancer cells, and that multiple malignant phenotypes exist. Also, the behavior of highly metastatic Calu-1 cells is sufficiently distinct from that of normal L-132 cells when placed on purified collagen IV to develop an accurate and quantitative model predicting metastatic potential.

3) Epigenetic silencing associated with aberrant methylation of promoter region CpG islands is one mechanism leading to loss of tumor suppressor function in human cancer. Using DNA pattern recognition we identified seven discriminatory novel patterns. When applying our discriminant analysis techniques, a classification function was derived based upon the frequency of these seven sequence patterns that was capable of discriminating methylation-prone from methylation-resistant CpG islands with 82% intrinsic susceptibility to de novo methylation, and suggest that the propensity for a CpG island to become aberrantly methylated can be predicted based on its sequence context.